

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 00:49:37 ; Search time 2942.31 Seconds  
(without alignments)  
13738.812 Million cell updates/sec

Title: US-09-768-781-2  
Perfect score: 1389  
Sequence: 1 atgaacacagaccaca.....caaggcaagtgtgtctga 1389

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb\_ba.\*
- 2: gb\_btg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_ey.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_ey.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	748.8	53.9	37650	9	HSU131B10	Z73417 Human DNA s
c 2	622.4	44.8	168918	2	AC116256	AC116256 Rattus no
c 3	617	44.4	171014	10	AL671915	AL671915 Mouse DNA
4	452.4	32.6	1609	9	AK098608	AK098608 Homo sapi
5	303.4	21.8	2911	10	AF155511	AF155511 Mus muscu
6	291.2	21.0	5096	6	AX332514	AX332514 Sequence
7	291.2	21.0	5096	9	HSXKWP	Z33684 Homo sapien
8	283.6	20.4	5215	6	AX405691	AX405691 Sequence
c 9	278	20.0	113688	9	AC005301	AC005301 Homo sapi
c 10	278	20.0	124823	9	AC007064	AC007064 Homo sapi
c 11	249.8	18.0	168918	2	AC116256	AC116256 Rattus no
c 12	218	15.7	160200	10	AC091605	AC091605 Mus muscu
c 13	218	15.7	200698	2	AL672060	AL672060 Mus muscu
c 14	217.2	15.6	16436	9	AC016752	AC016752 Homo sapi
15	215.6	15.5	161879	9	AC017005	AC017005 Homo sapi
16	215.6	15.5	182083	9	AC007965	AC007965 Homo sapi
17	210.8	15.2	37566	9	AC068541	AC068541 Homo sapi
c 18	210.8	15.2	174082	9	AC007379	AC007379 Homo sapi
c 19	207.6	14.9	101912	9	AC000100	AC000100 Homo sapi
c 20	207.6	14.9	120951	9	AC073649	AC073649 Homo sapi
21	207.6	14.9	145383	9	AC010682	AC010682 Homo sapi
c 22	207.6	14.9	149794	2	AC007273	AC007273 Homo sapi
23	194.4	14.0	175531	9	HSXKSRPXM	AL121577 Homo sapi
24	179.8	12.9	163495	9	AC024183	AC024183 Homo sapi
c 25	179.8	12.9	164891	2	AC022848	AC022848 Homo sapi
c 26	179.8	12.9	199174	9	AC007742	AC007742 Homo sapi
27	173	12.5	1588	9	HSFA000997	AF000997 Homo sapi
c 28	164	11.6	91640	2	AC123295	AC123295 Rattus no
29	161.8	11.6	626	6	AX079777	AX079777 Sequence
30	104.8	7.5	321	11	G03706	G03706 human STS W
31	91.6	6.6	164891	2	AC022848	AC022848 Homo sapi
c 32	62	4.5	64450	9	AC073893	AC073893 Homo sapi
c 33	57.8	4.2	138674	2	AC112084	AC112084 Rattus no
34	51.6	3.7	125020	9	AF429315	AF429315 Homo sapi
35	49.8	3.6	188048	10	AL669974	AL669974 Mouse DNA
36	47.4	3.4	212374	2	AC022773	AC022773 Mus muscu
c 37	46.8	3.4	68873	2	AC101531	AC101531 Mus muscu
c 38	46.8	3.4	234491	2	AL671118	AL671118 Mus muscu
c 39	45.6	3.3	144794	9	AC107055	AC107055 Homo sapi
c 40	45.4	3.3	7218	6	I66494	I66494 Sequence 14
c 41	45	3.2	1694	9	AF226053	AF226053 Homo sapi
42	44.4	3.2	13427	6	AX346828	AX346828 Sequence
43	44	3.2	152434	2	AC016262	AC016262 Homo sapi
44	43.8	3.2	104455	2	AC098390	AC098390 Rattus no
45	43.2	3.1	12229	3	AE001377	AE001377 Plasmodiu

ALIGNMENTS

RESULT 1	HSU131B10	DNA	linear	PRI 23-NOV-1999
LOCUS	Human DNA sequence from cosmid U131B10, between markers DXS366 and DXS87 on chromosome X contains XK membrane transport protein, ESTs and STS.	37650 bp		
DEFINITION	Human DNA sequence from cosmid U131B10, between markers DXS366 and DXS87 on chromosome X contains XK membrane transport protein, ESTs and STS.			
ACCESSION	Z73417			
VERSION	Z73417.1	GI:1322397		
KEYWORDS	X; XK membrane transport protein.			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 37650) Dodsworth, S.			

**TITLE** Direct Submission  
**JOURNAL** Submitted (17-MAY-1995) Sanger Centre, Hinxton, Cambridgeshire, CB10 1RQ, UK. E-mail enquiries: humquery@sanger.ac.uk  
**COMMENT** requests: clonerequest@sanger.ac.uk  
 IMPORTANT: This sequence is the entire insert of clone U131B10. The true left end of clone U131B10 is at 1 in this sequence. The true right end of clone U131B10 is at 37650.  
**FEATURES** U131B10 is from the human chromosome X-specific cosmid library.

source	Location/Qualifiers
1..37650	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="X"
	/map="X"
	/clone_lib="LLOXNC01"
364..519	/note="LLOXNC01-131B10"
repeat_region	364..519
repeat_region	/note="MIR element fragment"
2000..2290	/partial
repeat_region	/note="Alu repeat: matches 308..1 of consensus"
3347..3637	/note="Alu repeat: matches 1..308 of consensus"
repeat_region	3699..3800
repeat_region	/note="MSTA element fragment"
3712..3842	/note="THE1B element fragment"
repeat_region	3875..4004
repeat_region	/note="MSTA element fragment"
3909..4009	/note="THE1B element fragment"
repeat_region	4038..4085
repeat_region	/note="24 copies of 2 mer 85 & conserved"
repeat_region	4087..4378
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 308..1 of consensus"
misc_feature	4957..5561
misc_feature	/note="match: Multiple ESTs"
repeat_region	complement(5312..5561)
repeat_region	/note="match: STS G15323"
repeat_region	5723..5842
repeat_region	/note="L1 element fragment"
repeat_region	5980..6292
repeat_region	/note="MLR2C2 element fragment"
repeat_region	5980..6048
repeat_region	/note="MLR2D element fragment"
repeat_region	6399..6759
repeat_region	/note="L1 element fragment"
repeat_region	7148..7437
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 308..1 of consensus"
repeat_region	7460..7754
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 308..1 of consensus"
repeat_region	9496..9786
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 308..1 of consensus"
repeat_region	10339..10630
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 308..1 of consensus"
repeat_region	12989..13148
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 308..145 of consensus"
repeat_region	13166..13450
repeat_region	/note="Alu repeat: matches 1..308 of consensus"
repeat_region	14160..14208
repeat_region	/note="L1 element fragment"
repeat_region	16486..16792
repeat_region	/note="Alu repeat: matches 1..308 of consensus"
repeat_region	17220..17477
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 28..308 of consensus"
repeat_region	17909..18098
repeat_region	/note="Alu repeat: matches 308..1 of consensus"
repeat_region	32221..32499
repeat_region	/note="Alu repeat: matches 1..301 of consensus"
repeat_region	32500..32529
repeat_region	/note="15 copies of 2 mer 100 & conserved"
repeat_region	33138..33433
repeat_region	/note="Alu repeat: matches 1..308 of consensus"
repeat_region	33557..33894
repeat_region	/note="MLTID element fragment"
repeat_region	33914..33989
repeat_region	/note="MLTIE element fragment"
repeat_region	33935..34011
repeat_region	/note="MER3 element fragment"
repeat_region	18505..18795
repeat_region	/note="Alu repeat: matches 1..308 of consensus"
repeat_region	19300..19436
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 1..149 of consensus"
repeat_region	19440..19644
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 1..224 of consensus"
repeat_region	19645..19710
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 242..308 of consensus"
repeat_region	19711..19870
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 149..308 of consensus"
repeat_region	21881..22027
repeat_region	/note="MIR element fragment"
repeat_region	23139..23303
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 178..1 of consensus"
repeat_region	23343..23508
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 308..138 of consensus"
repeat_region	23509..23800
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 308..1 of consensus"
repeat_region	23840..23925
repeat_region	/note="MIR2 element fragment"
repeat_region	24408..24890
repeat_region	/note="L1 element fragment"
repeat_region	24527..24554
repeat_region	/note="14 copies of 2 mer 100 & conserved"
repeat_region	24902..24947
repeat_region	/note="23 copies of 2 mer 89 & conserved"
repeat_region	24951..25431
repeat_region	/note="L1 element fragment"
repeat_region	complement(25545..26171)
repeat_region	/partial
repeat_region	/note="match: 23684 XK membrane transport protein"
repeat_region	26196..26539
repeat_region	/note="match: 5' EST H87640 clone 252548; Paired with EST H87641 matching this cosmid"
repeat_region	complement(26828..27189)
repeat_region	/note="match: 3' EST H87629 clone 252524"
repeat_region	complement(26833..27186)
repeat_region	/note="match: 3' EST H87641 clone 252528; Paired with EST H87640 matching this cosmid"
repeat_region	27219..27266
repeat_region	/note="24 copies of 2 mer 96 & conserved"
repeat_region	27430..27721
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 308..1 of consensus"
repeat_region	28995..30018
repeat_region	/note="match: Multiple ESTs"
repeat_region	30042..30333
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 308..1 of consensus"
repeat_region	30691..30724
repeat_region	/note="17 copies of 2 mer 82 & conserved"
repeat_region	30725..31015
repeat_region	/partial
repeat_region	/note="Alu repeat: matches 308..1 of consensus"
repeat_region	32221..32499
repeat_region	/note="Alu repeat: matches 1..301 of consensus"
repeat_region	32500..32529
repeat_region	/note="15 copies of 2 mer 100 & conserved"
repeat_region	33138..33433
repeat_region	/note="Alu repeat: matches 1..308 of consensus"
repeat_region	33557..33894
repeat_region	/note="MLTID element fragment"
repeat_region	33914..33989
repeat_region	/note="MLTIE element fragment"
repeat_region	33935..34011

Qy	1326	TCGACACAGGGTTGAGAACTCAGAGCCACCCCTTTGAGACTGAGCAAGGCAAGGTTGTGT	1385
Db	26236	TCGACACAGGGTTGAGAACTCAGAGCCACCCCTTTGAGACTGAGCAAGGCAAGGTTGTGT	26295
Qy	1386	CTGA 1389	
Db	26296	CTGA 26299	
RESULT 2			
AC116256/c			
LOCUS	AC116256	168918 bp	DNA linear HTG 18-JUL-2002
DEFINITION	Rattus norvegicus clone CH230-139K19, *** SEQUENCING IN PROGRESS		
	***, 74 unordered pieces.		
ACCESSION	AC116256		
VERSION	168918		
KEYWORDS	HTG: HTGS PHASE1.		
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1 (bases 1 to 168918)		
AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oman,F.R., Allen,C., Alsbrooks,S.L., Amaratus,H.C., Are,J.R., Ayele,M., Banks,T., Barbieri,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D., Bouck,J.J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homst,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Krasov,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratochvic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseg,H., Lozada,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., Mcleod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scher,S., Scott,G., Shen,H., Shoohtani,N., Sison,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansay,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villafra,B., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.		
	Direct Submission		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 168918)		
AUTHORS	Worley,K.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
REFERENCE	3 (bases 1 to 168918)		
AUTHORS	Worley,K.C.		

TITLE  
JOURNAL

Direct Submission  
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

## COMMENT

On Jul 14, 2002 this sequence version replaced gi:19718450.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GTLM

Center clone name: CH230-139K19

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 104166 bases at least Q40

Consensus quality: 113489 bases at least Q30

Consensus quality: 119214 bases at least Q20

-----

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 74 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 1568: contig of 1568 bp in length  
\* 1569 1668: gap of unknown length  
\* 1669 2687: contig of 1019 bp in length  
\* 2688 2787: gap of unknown length  
\* 2788 4137: contig of 1350 bp in length  
\* 4138 4237: gap of unknown length  
\* 4238 5268: contig of 1031 bp in length  
\* 5269 5368: gap of unknown length  
\* 5369 6665: contig of 1297 bp in length  
\* 6666 8226: contig of 1461 bp in length  
\* 8227 8326: gap of unknown length  
\* 8327 9449: contig of 1123 bp in length  
\* 9450 9549: gap of unknown length  
\* 9550 10623: contig of 1074 bp in length  
\* 10624 10723: gap of unknown length  
\* 10724 11866: contig of 1043 bp in length  
\* 11867 12876: contig of 1010 bp in length  
\* 12877 12976: gap of unknown length  
\* 12977 14300: contig of 1324 bp in length  
\* 14301 14400: gap of unknown length  
\* 14401 15614: contig of 1214 bp in length  
\* 15615 15714: gap of unknown length  
\* 15715 17724: contig of 2010 bp in length  
\* 17725 17824: gap of unknown length  
\* 17825 19055: contig of 1231 bp in length  
\* 19056 19155: gap of unknown length  
\* 19156 20675: contig of 1520 bp in length  
\* 20676 20775: gap of unknown length  
\* 20776 22570: contig of 1795 bp in length  
\* 22571 22670: gap of unknown length  
\* 22671 24166: contig of 1496 bp in length  
\* 24167 24266: gap of unknown length  
\* 24267 25499: contig of 1233 bp in length  
\* 25500 25599: gap of unknown length  
\* 25600 27074: contig of 1475 bp in length  
\* 27075 27174: gap of unknown length  
\* 27175 28475: contig of 1301 bp in length  
\* 28476 28575: gap of unknown length  
\* 28576 29878: contig of 1303 bp in length  
\* 29879 29978: gap of unknown length

\* 29979 31798: contig of 1820 bp in length  
\* 31799 31998: gap of unknown length  
\* 31999 33456: contig of 1558 bp in length  
\* 33457 33556: gap of unknown length  
\* 33557 34753: contig of 1197 bp in length  
\* 34754 34853: gap of unknown length  
\* 34854 37077: contig of 2224 bp in length  
\* 37078 37177: gap of unknown length  
\* 37178 38959: contig of 1782 bp in length  
\* 38960 39059: gap of unknown length  
\* 39060 40603: contig of 1544 bp in length  
\* 40604 40703: gap of unknown length  
\* 40704 41879: contig of 1176 bp in length  
\* 41880 41979: gap of unknown length  
\* 41980 44025: contig of 2046 bp in length  
\* 44026 44125: gap of unknown length  
\* 44126 45351: contig of 1226 bp in length  
\* 45352 45451: gap of unknown length  
\* 45452 47009: contig of 1558 bp in length  
\* 47010 47109: gap of unknown length  
\* 47110 49436: contig of 2327 bp in length  
\* 49437 49536: gap of unknown length  
\* 49537 51265: contig of 1729 bp in length  
\* 51266 51365: gap of unknown length  
\* 51366 52885: contig of 1520 bp in length  
\* 52886 52985: gap of unknown length  
\* 52986 55166: contig of 2181 bp in length  
\* 55167 55266: gap of unknown length  
\* 55267 57181: contig of 1915 bp in length  
\* 57182 57281: gap of unknown length  
\* 57282 59138: contig of 1857 bp in length  
\* 59139 59238: gap of unknown length  
\* 59239 61142: contig of 1904 bp in length  
\* 61143 61242: gap of unknown length  
\* 61243 63220: contig of 1978 bp in length  
\* 63221 63320: gap of unknown length  
\* 63321 64522: contig of 1202 bp in length  
\* 64523 64622: gap of unknown length  
\* 64623 66069: contig of 1447 bp in length  
\* 66070 66169: gap of unknown length  
\* 66170 67304: contig of 1135 bp in length  
\* 67305 67404: gap of unknown length  
\* 67405 69305: contig of 1901 bp in length  
\* 69306 69405: gap of unknown length  
\* 69406 71453: contig of 2048 bp in length  
\* 71454 71554: gap of unknown length  
\* 71554 74540: contig of 2987 bp in length  
\* 74541 74640: gap of unknown length  
\* 74641 76351: contig of 1711 bp in length  
\* 76352 76451: gap of unknown length  
\* 76452 78243: contig of 2792 bp in length  
\* 78244 79343: gap of unknown length  
\* 79344 81674: contig of 2331 bp in length  
\* 81675 81774: gap of unknown length  
\* 81775 83496: contig of 1722 bp in length  
\* 83497 83596: gap of unknown length  
\* 83597 85142: contig of 1546 bp in length  
\* 85143 85242: gap of unknown length  
\* 85243 87769: contig of 2527 bp in length  
\* 87770 87869: gap of unknown length  
\* 87870 89602: contig of 1733 bp in length  
\* 89603 89702: gap of unknown length  
\* 89703 91131: contig of 1429 bp in length  
\* 91132 91231: gap of unknown length  
\* 91232 93807: contig of 2576 bp in length

Query Match 44.8%; Score 622.4; DB 2; Length 168918;

Best Local Similarity 98.6%; Pred. No. 1.9e-167;

Matches 674; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

OY 629 TTCCTTATTTGCCAGCTGTGTAATGCTATTTTCCTGGTATCTGTACCTATGGGGCCA 688

Db 140555 TTCCTTATTTGCCAGCTGTGTAATGCTATTTTCCTGGTATCTGTACCTATGGGGCTA 140496







D	591	GTGGAAGAAGTCTCCTCATGACCATATCCCTGTTGTCATATGTTATGAGCCTTGCGC	640
Q	694	TGCAATATGTTGGCTATCCAGATCAAGTACGATACAGATTCGCCTTGGGCCACTA	753
D	641	TGCAACATCCTAGCCATCAAAATCAAGTACGATGATGAGTCAAAGTGAAGCCTCTG	700
Q	754	GAAGTCTCTGCATCAGCATCTGGCGGACATGGAGATCACTTCCCGCCTCTGATTCTG	813
D	701	GCTATGTCGTATCTTCTCTGTGGAGGAGCTTTGAGATTGCCATTCGAGTTGATGCTG	760
Q	814	GTGCTCTTCTCAGCCACTTTGAAATTGAAGCTGTGCCCTTCTTAGTGCTCAACTTCCTG	873
D	761	GTCTCTTTACTCCGCTCGAAGACCTGGGTGGTTATATATCATCAACTTCTTC	820
Q	874	ATCATCTCTTTGAGCCCTCGGATTAAGTTCTTGGAGAAAGTGTGCCAGATGCCCAATAAC	933
D	821	AGTTTCTTTGTACCCCTGGATCCTCTTCTGTGTGACGTGTTCCCCATTCCCTGAGAAC	880
Q	934	ATTGAGAAAAAATTCACGCGGGTGGCACTCTGTGGTCTGTGATTTTCAGTCACCATCCTC	993
D	881	ATAGAGAAGGCCCTCAGTAGAGTGGGCACCAACCATTTGTACTATGCTTTTCTAACTTTACTC	940
Q	994	TATGCTGCATCAACTTCTCTGTGTGTGACGTTTGAGTTTGAGGTTGGCAGACAGAGAT	1053
D	941	TATACTGTATCAACATGTTCTGTGTGTCTGCTGTACAGCTGAAATTTGACAGCCCTGAC	1000
Q	1054	CTGCTCGACAAAGGGCAGAACTGGGGACATATGGGCTGCACATATAGTGTGAGGTTGGTA	1113
D	1001	CTCATCAGCAAGTCCCATATTTGGTACCACTACTGGTGTATTACATGATAAGATTTCATC	1060
Q	1114	GAGAAATGTGATCATGCTCTTGGTTTTTAAGTTCTTTTGGAGTGAAAGTGTTCAGTAATAC	1173
D	1061	GAGAAATGCCATCTCTCTCTCTGCTGCTGCTGTATCTTTTCAAGACTGCACATCTATATGATGTG	1120
Q	1174	TGTCATTCCTTGATGCTTGGCAGCTATTAATTGCTTATCTGATTTCCTATGGCTTCATG	1233
D	1121	TGCGACCTCTGTTGGTCTGTCAGCTGCTCATTTGGGTAATGTCAGACGCAATCTCTTCATG	1180
Q	1234	CTCCTTTTCTTCCAGTACTTGCATCCATTTGGCTCACTCTTACCACATATGTAGTAGAC	1293
D	1181	CTTGATTCTATCAGTTCTTCCACCTTGCAAAAGCTCTTTTCTTCCAGTGTTCTGAA	1240
Q	1294	TACCTCCA	1301
D	1241	GGCTTTCA	1248
RESULT	7		
LOCUS	HSXKMTP	5096 bp	linear
DEFINITION	Homo sapiens mRNA for membrane transport protein (XK gene).		
ACCESSION	232684		
VERSION	232684.2	GI:4883432	
KEYWORDS	membrane transport protein; XK gene.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Ho, M., Chelly, J., Carter, N., Dane, A., Crocker, P. and Monaco, A. P.		
TITLE	Isolation of the gene for McLeod syndrome that encodes a novel membrane transport protein		
JOURNAL	Cell 77 (6), 869-880 (1994)		
MEDLINE	94273191		
PUBMED	8004674		
REFERENCE	2 (bases 1 to 5096)		
AUTHORS	Ho, M. F.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-APR-1994) Meng F Ho, Human Genetics, Imperial Cancer Research Fund, Institute of Molecular Medicine John Radcliffe Hospital Headington, Oxford, OX3 9DU, United Kingdom		
REMARK	revised by [3]		
REFERENCE	3 (bases 1 to 5096)		

## AUTHORS

HO, M.F.

## TITLE

Direct Submission  
Submitted (21-OCT-1997) Meng P Ho, Human Genetics, Imperial Cancer  
Research Fund, Institute of Molecular Medicine John Radcliffe  
Hospital Headington, Oxford, OXON, OX3 9DU, United Kingdom  
On May 20, 1999 this sequence version replaced gi:2570027.

## COMMENT

Location/Qualifiers

## source

1. 5096

/organism="Homo sapiens"

/db xref="taxon:9606"

/chromosome="X"

/clone="XK"

83. .5068

/gene="XK"

83. .1417

/gene="XK"

/citation=[1]

/codon\_start=1

/product="membrane transport protein"

/protein\_id="CAA83632.2"

/db xref="GI:4883433"

/translation="MKPPASVLAFLVFAETTAALSLSTYSBGGDMQWQALTLLES  
LLPCALVQLTLLVHRDLSDRLPLVLLHLQLPLRPFVFCIIPQSGNNEPIVS  
ITKKQPKNGLSSEEIEKEVQAGKLIHRSAPRASVIOAFLGAPQUTLQIYISV  
MQQDVTVGRSLMTISLLSIVYGALRNILAIKIYDEYEVKVKPLAYVICIFLWRGFE  
IATRVVVLFTSLVLTWVVIILINPFSFLPWLFWCGSGSPFENIEKALSRVGT  
TIVLCFLTLTYTGNNFCWSAVOLKIDSPDLISKSHNWYQLLYVYIRFIENAILLLL  
WLPKTDIYNYVCAPLLVLQLLYGYCTAILFLVYQFPFKKILFSSSVSEGFQRLW  
RCFCWACROQKCEPIGKEDLQSDRDRDETFSSSKTSPFPGFLNAEDLCSN"

83. .327

/gene="XK"

/number=1

328. .590

/gene="XK"

/number=2

591. .1417

/gene="XK"

/number=3

5036. .5041

/gene="XK"

5063. .5068

/gene="XK"

/polya\_site

5092

BASE COUNT 1392 a 1064 c 1022 g 1618 t

## ORIGIN

Query Match 21.0%; Score 291.2; DB 9; Length 5096;

Best Local Similarity 54.7%; Pred. No. 3e-72;

Matches 628; Conservative 0; Mismatches 508; Indels 12; Gaps 2;

Qy 157 TCCACCTTTTGTACTGTGGGAGGCTGCATCTGCTTTGTACATGGTTAGNAATCTATCGA 216

Db 110 TCGTGTCTCTGTTCGTGGCGGACACACGGCGCGCTGACGCTGAGCAGCACCTACCGC 169

Qy 217 AAGAAATAGTGAACCTTACTGGATGACATACACCTTTTCTTTTATGTTTTCATCCATT 276

Db 170 TCGGGCGGACCGCATGTGGCAGCGCGCTGACGTTGCTTTCTCGCTACTCCCTTGGCG 229

Qy 277 ATGGTCCAGTTGACCTTCATTTTGTCCACAGAGATCTAGCCAAAGATAAACCGCTATCA 336

Db 230 CTGTCGAGCTCAGCTTCTCTTCGTACACCGGACCTCAGCGCGCGCGCTCGTA 289

Qy 337 TTATTTATGATCATCTTCTTGGGACCTGTTATCAGATGTTTGGAGGCGCATGATTAAG 396

Db 290 CTGCTGTCGACCTGCTGCACTTTGGGCGCCCTTTTCAAGTGTGTTTGAAGTCTTCTGCA 349

Qy 397 TACCTCACACTGTGGAGAAGAGGACGAGGAGCGCCCTATGTCAGCCTCACCCGAAG 456

Db 350 TACTTTC-----AGTCAGGCAACAATGAAGAGCCTTATGTCAGTATCACCAAGAAG 400

Qy 457 AAG---ATGCTAATAGATGGCAGGAGGTGCTGTAGTAATGGGAGGTGGGCGACTCCATC 513

Db 401 AGGCAATGCCAAAATAGGCTCTCAGAGGAGATTGAGAGGAGGTGGGCGGCGAGAA 460

Qy 514 CGGACCTTGGCTATGTCACCGCAATGSCCTACAAAAGTATGTTGACAGATCCAAAGCCTTCTTG 573

Db 461 GGCATACTAATACCCACCGATCAGCTTACCGGGGCTGGTATCCAGGCTTTCTTG 520

Qy 574 GGCTCAGTGGCCCCAGCTGACCTATCAGCTTATGTGAGCCCTGATCTCTGACAGGTTCCC 633

Db 521 GGCTCAGCCCCCAGCTGACCTACAGCTGTACATAAGTGTCTATGAGGAGCGTCACT 580

Qy 634 CTGGGTAGAGTTGTCTAATGGTATTTTCCCTGGTATCTGTACCTATGGGGCCACCTT 593

Db 581 GTTGAAGAAGTCTCTCTCATGACCATATCCCTGTTGTCCATTTGTATGGAGCCTTGGCG 640

Qy 694 TGCAATATGTTGGCTATCCAGATCAAGTACGATGACTCAAGATTGCGCTTGGGCCACTA 753

Db 641 TGCAATATCTAGCCATCAAAATCAAGTACGATGATGATGAGTCAAGTGAAGCCTCTG 700

Qy 754 GAAGTCTCTGCAATCAGCATCTGGCGGACATTTGGAGATCATTCCCGCTCTCTGATTCG 813

Db 701 GCCTATGCTCTATCTCTCTGAGGAGCTTTGAGATTGGCAGCTCGAGTTGTAGTCTTG 760

Qy 814 GTGCTCTTCTCAGCAGCTTTGAAATGAGGCTGTGCCCTTCCCTAGTGTCTCACTTCTTG 873

Db 761 GTCTCTTTTACCTCGCTCTGAGACCTGGGTGGTGGTTATAATACTCATCAACTTCTTC 820

Qy 874 ATCATCTCTTTGAGCCCTGGAATTAAGTTCTGGAGAAGTGGTGGCCAGATGCCAATAAC 933

Db 821 AGTTCTTTTGTATACCTCTGATCTCTTCTGGTGCAGTGGTTCCTCCATTTCCCTGAGA 880

Qy 934 ATTGAGAAAACTTTGAGCCGGCTGCGCACTCTGGTGGTCTCTGATTTCACTCACTTCTC 993

Db 881 ATAGAGAAGGCTCTGATGAGTGGCAGCACCATTGTACTATGCTTTCTAATTTACTTC 940

Qy 994 TATCTGGATCAACTCTCTTCTGCTGGTCACTTTGCTAGTTGAGGTTGGCAGACAGAT 1053

Db 941 TATCTGGTATCAACATGTTCTGCTGGTCTCTGTACAGCTGAAATTTGACAGCCCTGAC 1000

Qy 1054 CTGCTGACAAAGGCGAGACTGGGGACATATGGCCCTGCACTATAGTGTGAGGTTGTA 1113

Db 1001 CTATCAGAAAGTCCCAATATTTGTTACAGCTCTGTTTCAAGACTGACATCTATATGATG 1060

Qy 1114 GAGAATGTGATCATGCTCTGGTGTGTTTAAAGTCTTTTGGAGTCAAAAGTGTACTGAATTAC 1173

Db 1061 GAGAATGCCATCT 1120

Qy 1174 TGTCACTCTTGTGATGCTTGGCAGCTCATTTATGCTTATCTGATTTTCCATTTGGCTTCA 1233

Db 1121 TGGCAGCTCTGTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1180

Qy 1234 CTCTTTTCTTCCAGTACTTGGCTCCTTGGCTGCTCTTCCACCCATATGCTAGTAGAC 1293

Db 1181 CTGTGATTTCTATCAGTCTTCTCCACCTTTGCAAAAGCTCTTTTCTTCCAGTGTCTTCT 1240

Qy 1294 TACCTCCA 1301

Db 1241 GGCTTTCA 1248

## RESULT 8

AX405691

LOCUS

DEFINITION Sequence 106 from Patent WO0222660.

AX405691

ACCESSION

VERSION AX405691.1 GI:21438821

KEYWORDS

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Theria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS

Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,

Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.

TITLE Novel nucleic acids and polypeptides

AX405691 5215 bp DNA linear PAT 14-JUN-2002



there are numerous instances of insertions, deletions, and single nucleotide polymorphisms in the overlapping regions below.  
AC005301(p15j16) 77414 113688 (0) overlaps AC007064(p8708) 1 36275 (88548).

```

FEATURES             Location/Qualifiers
     source            1..113688
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="22q11"
                        /clone="p15j16"
BASE COUNT           37158 a 22385 c 21989 g 32256 t
ORIGIN

Query Match          20.0%; Score 278; DB 9; Length 113688;
Best Local Similarity 65.1%; Pred. No. 3.1e-68;
Matches 410; Conservative 0; Mismatches 220; Indels 0; Gaps 0;

QY 640 AGAGTTGCTAATGGTATTTCCCTGGTATCTGTACCTATGGGGCCACCTTTGGCAAT 699
Db 101529 ATAGCAATTGCTGATGACATATTTCCCTGTTATCAGTTACTTATGGGGCCATTCGCTGCAAT 101470

QY 700 ATGTGGCTATCCAGATCAAGTACGATGACTACAAAGATTTCGCTTTGGGCCCACTAGAGTC 759
Db 101469 ATACTGGCCATCCAGATCAGCAATGATGATCTACCAITTAAGCTACCGCCGATAGAATTC 101410

QY 760 CTCTGCATCACCATCTGGCGGACATGGAGATCACTTCCCGCTCTCTGATTCCTGGTGTC 819
Db 101409 TCTGTGTCGTGATGGGGTTTTTGGAGGTATCTACGTTAGTACTCTGCAATTT 101350

QY 820 TTCTCAGCCATTTGAAATGAAGCTGTGCCCTTCTTAGTGCTCAACTTCTCGATCATC 879
Db 101349 TTCATTGCATCTCTGAACTGAAGAGCTACCCGTTTGTGTAATCATATATTTTGTATCA 101290

QY 880 CTCCTTGAGCCCTGATTAAGTCTCGAGAAGTGTGGCCAGATGCCCAATTAACATTGAG 939
Db 101289 TTGTTGGCACCCTGGCTGGAGTTTGGAAAAGTGGAGCTCATCTTCTCGCAACAAGAA 101230

QY 940 AAAAATTCAGCCGGTGGCACTCTGTGGTCTGATTTTCAGTCACCATCTCTATGCT 999
Db 101229 ATAATTCATATATGGTGGTACAGTACTGATCTTTCTTGATCACAAGTATATGCT 101170

QY 1000 GGCATCAACTTCTTCTGTGCTGAGCTTTGCAGTTGAGTTGGCAGACAGAGATCTCGTC 1059
Db 101169 GCCATCAACTTCTCTGCTGTGTCAGCACTGAACTGCAGTTGTGAGTACAAATATTT 101110

QY 1060 GACAAAGGCGAAGTGGGACATATGGCCCTGCACTATAGTGTGAGTTGTGAGAGAT 1119
Db 101109 GACGGGAGACAGAGTGGGGCCATAGAAATCCTACATCAGCTTTTCAGTTTTTAGAAAAT 101050

QY 1120 GTGATCATGCTTGTGTTTAAAGTTCTTTGGAGTGAAGTTTACTGAATTAAGTCTCAT 1179
Db 101049 GTGATTAATGATATTTGTTATTTAGGTTCTTTGGAGGAAACCTTTGCTGAATTTGTGAC 100990

QY 1180 TCCTTGATGCTTGCAGCTCAATTATGCTTATCTGATTTTCATTTGCTTTCATGCTCCTT 1239
Db 100989 TCATTAATGCGGTGACGCTCATATAGCTACCTATTGGCCACTGGCTTTATGCTCTC 100930

QY 1240 TTCTTCAGTACTTGCATCCATTTGGGCTCA 1269
Db 100929 TTCTATCAGTATTTGTACCCATGGCAGTCA 100900

RESULT 10
AC007064/c          124823 bp DNA linear PRI 27-MAY-2000
LOCUS               Homo sapiens Chromosome 22q11 PAC Clone p8708 in CES region,
DEFINITION          complete sequence.
ACCESSION           AC007064
VERSION              AC007064.27 GI:7109507
KEYWORDS             HTG.
SOURCE               Homo sapiens.
ORGANISM             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 124823)
Ray, L.A., Loh, P., Qi, S., Sloan, D., McDermid, H. and Roe, B.A.
Homo sapiens Chromosome 22q11 PAC Clone p8708 in CES Region
Unpublished
2 (bases 1 to 124823)
Ray, L.A., Loh, P., Qi, S., Sloan, D., McDermid, H. and Roe, B.A.
Direct Submission
Submitted (11-MAR-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 124823)
Ray, L.A., Loh, P., Qi, S., Sloan, D., McDermid, H. and Roe, B.A.
Direct Submission
Submitted (26-OCT-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 124823)
Ray, L.A., Loh, P., Qi, S., Sloan, D., McDermid, H. and Roe, B.A.
Direct Submission
Submitted (27-OCT-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
5 (bases 1 to 124823)
Ray, L.A., Loh, P., Qi, S., Sloan, D., McDermid, H. and Roe, B.A.
Direct Submission
Submitted (05-NOV-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
6 (bases 1 to 124823)
Ray, L.A., Loh, P., Qi, S., Sloan, D., McDermid, H. and Roe, B.A.
Direct Submission
Submitted (08-NOV-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
7 (bases 1 to 124823)
Ray, L.A., Loh, P., Qi, S., Sloan, D., McDermid, H. and Roe, B.A.
Direct Submission
Submitted (29-FEB-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
8 (bases 1 to 124823)
Ray, L.A., Loh, P., Qi, S., Sloan, D., McDermid, H. and Roe, B.A.
Direct Submission
Submitted (25-MAY-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
9 (bases 1 to 124823)
Ray, L.A., Loh, P., Qi, S., Sloan, D., McDermid, H. and Roe, B.A.
Direct Submission
Submitted (27-MAY-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Feb 29, 2000 this sequence version replaced gi:6249691.
Because these overlapping clones came from different libraries
there are numerous instances of insertions, deletions, and single
nucleotide polymorphisms in the overlapping regions below.
AC005301(p15j16) 77414 113688 (0) overlaps AC007064(p8708) 1 36275
(88548) AC007064(p8708) 62616 124823 (0) overlaps AC006548(p20k14)
1 62199 (137950).
FEATURES             Location/Qualifiers
     source            1..124823
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="22q11"
                        /clone="p8708"
BASE COUNT           41074 a 23932 c 23324 g 36493 t
ORIGIN

Query Match          20.0%; Score 278; DB 9; Length 124823;
Best Local Similarity 65.1%; Pred. No. 3.2e-68;
Matches 410; Conservative 0; Mismatches 220; Indels 0; Gaps 0;

```

```

QY 640 AGAGTTGGCTAATGGTATTTTCCTGGTATCTGTCACTATGGGGCCACCCTTTGCAAT 699
|||
Db 24116 ATAGCATTTGCTGATGACATATTTTCCCTGTTATCATCAGTTACTTATGGGGCCATCGCTGCAAT 24057
|||
QY 700 ATGTTGGCTATCCAGATCAAGTACATCAAGATTCGGCTTGGGGCCACTAGAAGTC 759
|||
Db 24056 ATACTGGCCATCCAGATCAGCAATGATGATACCTACCATTAAGCTACCCGCCGATAGAAATTC 23997
|||
QY 760 CTCTGTCATCACCCTCGCGGACATTTGAGATCACTTCCCGCTCTCTGATTTCTGGTGTC 819
|||
Db 23996 TTCTGTGCTGATGCTGGCGTTTCTTGGAGGTTATCTCAGCTGTAGTACCTTGCGCATTT 23997
|||
QY 820 TTCTCAGCCACTTTGAAATGAAGCTGTGCCCTTCTTAGTGTCTCAACTTCTCTGATCATC 879
|||
Db 23936 TTCAATGTCATCTCTGAACTGAAGAGCTACCCCGTTTGTGTAATCATATATTTTGTATCA 23877
|||
QY 880 CTCTTTGAGCCCTGGATTAAGTTCTGGAGAAAGTGTGGCCAGATGCCCAATAACATTAAG 939
|||
Db 23876 TTGTGTGGCACCCTGGCTGGAGGTTTGGAAAAAGTGGAGGCTCATCTTCTGGCAACAAGAA 23817
|||
QY 940 AAAAATTCAGCCGGTGGCACTCTGGTGGTCTGATTTTCAGTCACCATCTCTCTATGCT 999
|||
Db 23816 AATAATTTCCATATGTTGGGTACAGTACTGATGCTTTCTTGTGATCACACTGCTATATGCT 23757
|||
QY 1000 GGCATCAACTTCTCTTGTGCTGCTGAGCTTTGAGGTTGAGGTTGGCAGACAGAGATCTCGTC 1059
|||
Db 23756 GCCATCAACTTCTCTGCTGCTGAGCTGAACTGCACTGTTGTCAGATGACAAATAATTT 23697
|||
QY 1060 GACAAAGGCGAGAACTGGGGACATATGGGCTGCACTATAGTGTGAGGTTGGTAGAGAAAT 1119
|||
Db 23696 GACGGGACAGAGATGGGGCCATAGAAATCTTACACTACAGCTTTCAGTTTTCAGTAAAT 23637
|||
QY 1120 GTGATCATGCTGCTGTTTAACTTCTTGGAGTGAAGTGTACTGATGCTTCTGATGATGCTAT 1179
|||
Db 23636 GTGATATGATATTTGTTATTTAGGTTCTTGGAGGGAATCTTCTGATGATGCTGATGCT 23577
|||
QY 1180 TCCTTTGCTGCTTGGCAGCTATTTGCTGCTATCTGATTTCCATGCTTTCATGCTCTCTT 1239
|||
Db 23576 TCATTAATGGCGTGCAGCTCATCATAGCTACCTATTGGCCACTGGCTTATGCTCTCTC 23517
|||
QY 1240 TTCTTCCAGTACTTGCATPCATTTGGCTCA 1269
|||
Db 23516 TTCTATCAGTATTTGTACCCATGGCAGTCA 23487
|||

RESULT 11
AC116256
LOCUS AC116256
DEFINITION Rattus norvegicus clone CH230-139K19, *** SEQUENCING IN PROGRESS
***, 74 unordered pieces.
AC116256
VERSION AC116256.2 GI:21745900
KEYWORDS HTG; HTGS PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 168918)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Albrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Deigado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

```

```

Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseghe,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G.,
Oggunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Roife,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonaika,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinsan,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 168918)
Worley,K.C.
Direct Submission
Submitted (26-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 168918)
Worley,K.C.
Direct Submission
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:19718450.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GTLM
Center clone name: CH230-139K19
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 104166 bases at least Q40
Consensus quality: 113489 bases at least Q30
Consensus quality: 119214 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 74 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1568: contig of 1568 bp in length
* 1569 1668: gap of unknown length
* 1669 2687: contig of 1019 bp in length
* 2688 2787: gap of unknown length
* 2788 4137: contig of 1350 bp in length
* 4138 4237: gap of unknown length

```

\* 4238 5268: contig of 1031 bp in length  
\* 5368: gap of unknown length  
\* 5369 6665: contig of 1297 bp in length  
\* 6666 6665: gap of unknown length  
\* 6766 8226: contig of 1461 bp in length  
\* 8227 8328: gap of unknown length  
\* 8327 9449: contig of 1123 bp in length  
\* 9450 9459: gap of unknown length  
\* 9550 10623: contig of 1074 bp in length  
\* 10624 10723: gap of unknown length  
\* 10724 11766: contig of 1043 bp in length  
\* 11767 11866: gap of unknown length  
\* 11867 12876: contig of 1010 bp in length  
\* 12877 12976: gap of unknown length  
\* 12977 14300: contig of 1324 bp in length  
\* 14301 14400: gap of unknown length  
\* 14401 15614: contig of 1214 bp in length  
\* 15615 15714: gap of unknown length  
\* 15715 17724: contig of 2010 bp in length  
\* 17725 17824: gap of unknown length  
\* 17825 19085: contig of 1231 bp in length  
\* 19086 19155: gap of unknown length  
\* 19156 20675: contig of 1520 bp in length  
\* 20676 20775: gap of unknown length  
\* 20776 22570: contig of 1795 bp in length  
\* 22571 22670: gap of unknown length  
\* 22671 24166: contig of 1496 bp in length  
\* 24167 24266: gap of unknown length  
\* 24267 25499: contig of 1233 bp in length  
\* 25500 25599: gap of unknown length  
\* 25600 27074: contig of 1475 bp in length  
\* 27075 27174: gap of unknown length  
\* 27175 28475: contig of 1301 bp in length  
\* 28476 28575: gap of unknown length  
\* 28576 29878: contig of 1303 bp in length  
\* 29879 29978: gap of unknown length  
\* 29979 31798: contig of 1820 bp in length  
\* 31799 33456: gap of unknown length  
\* 33457 33557: gap of unknown length  
\* 34754 34853: gap of unknown length  
\* 34854 37077: contig of 2224 bp in length  
\* 37078 37177: gap of unknown length  
\* 37178 38959: contig of 1782 bp in length  
\* 39059 39059: gap of unknown length  
\* 39600 40603: contig of 1544 bp in length  
\* 40604 40703: gap of unknown length  
\* 40704 41879: contig of 1176 bp in length  
\* 41880 41979: gap of unknown length  
\* 41980 44025: contig of 2046 bp in length  
\* 44026 44125: gap of unknown length  
\* 44126 45351: contig of 1226 bp in length  
\* 45352 45451: gap of unknown length  
\* 45452 47009: contig of 1558 bp in length  
\* 47010 47109: gap of unknown length  
\* 47110 49436: contig of 2327 bp in length  
\* 49437 49536: gap of unknown length  
\* 49537 51265: contig of 1729 bp in length  
\* 51266 51365: gap of unknown length  
\* 51366 52885: contig of 1520 bp in length  
\* 52886 52985: gap of unknown length  
\* 52986 55166: contig of 2181 bp in length  
\* 55167 55266: gap of unknown length  
\* 55267 57181: contig of 1915 bp in length  
\* 57182 59138: contig of 1857 bp in length  
\* 59139 59238: gap of unknown length  
\* 59239 61142: contig of 1904 bp in length  
\* 61143 61242: gap of unknown length  
\* 61243 63220: contig of 1978 bp in length  
\* 63221 63320: gap of unknown length  
\* 63321 64522: contig of 1202 bp in length

\* 64523 64622: gap of unknown length  
\* 64623 66069: contig of 1447 bp in length  
\* 66070 66169: gap of unknown length  
\* 66170 67304: contig of 1135 bp in length  
\* 67305 67405: gap of unknown length  
\* 67406 69305: contig of 1901 bp in length  
\* 69306 69405: gap of unknown length  
\* 71453 71553: contig of 2048 bp in length  
\* 71554 74540: contig of 2987 bp in length  
\* 74541 74640: gap of unknown length  
\* 74641 76351: contig of 1711 bp in length  
\* 76352 79243: contig of 2792 bp in length  
\* 79244 79343: gap of unknown length  
\* 79344 81674: contig of 2331 bp in length  
\* 81675 81775: gap of unknown length  
\* 81776 83596: gap of unknown length  
\* 83597 85142: contig of 1546 bp in length  
\* 85143 85242: gap of unknown length  
\* 85243 87769: contig of 2527 bp in length  
\* 87770 87869: gap of unknown length  
\* 87870 89602: contig of 1733 bp in length  
\* 89603 91131: contig of 1429 bp in length  
\* 91132 91231: gap of unknown length  
\* 91232 93807: contig of 2576 bp in length

Query Match 18.0%; Score 249.8; DB 2; Length 168918;  
Best Local Similarity 93.9%; Pred. No. 4.5e-60;  
Matches 260; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 372 CAGATGTTGGAGGCCATGATTAAGTACCTCACACTGTGGAGAAAGAGGACGAGGA 431  
Db 122642 CAGATGTTGGAGGCCATGATTAAGTACCTCACACTGTGGAGAAAGAGGACGAGGA 122701  
Qy 432 GCCCTATGTCAGCCTCACCCGAAAGAGATGCTAATAGATGCGAGGAGGTCCTGATAGA 491  
Db 122702 GCCATATGTCAGCCTCACCCGAAAGAGATGCTAATAGATGCGAGGAGGTCCTGATAGA 122761  
Qy 492 ATGGGAGGTGGGCCATCCATCCGACCCCTGGCTATGCACCGCAATGCTACAAACGTAT 551  
Db 122762 ATGGGAGGTGGGCCATCCATCCGACCCCTGGCTATGCACCGCAATGCTACAAACGTAT 122821  
Qy 552 GTCACAGATCCAGCCTTCCTGGGCTAGTCCGCCAGCTGACCTATACGCTTCTATGTGAG 611  
Db 122822 GTCACAGATCCAGCCTTCCTGGGCTAGTCCGCCAGCTGACCTATACGCTTCTATGTGAG 122881  
Qy 612 CCTGATCTCTGCAGAGTTCCTGGGTAGAGTTGTG 648  
Db 122882 CCTGATCTCTGCAGAGTTCCTGGGTAGAGTTGTG 122918

RESULT 12  
AC091605/c  
LOCUS AC091605 160200 bp DNA linear ROD 02-AUG-2002  
DEFINITION Mus musculus Strain C57BL6/J Chromosome X BAC, RP23-64A09, Complete  
Sequence, complete sequence.  
ACCESSION AC091605  
VERSION AC091605.6 GI:22091329  
KEYWORDS HTG.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 160200)  
AUTHORS Montgomery, K.T., Grills, G., Li, L., Chiu, D., Decker, J., Fusina, M.,  
Goltz, J., Haider, A., Hall, L., Han, J., Ioshikhes, I.P., Lee, E.,  
Perera, A., Shim, C., Thomas, E. and Kucherlapati, R.  
TITLE High Throughput Mouse Sequencing  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 160200)



Db	142978	CTGTTTCAGCTGAJAATCGACAAATCCGGAGCTCATCAGCAAGTCCAGAACTCGTACCGCC	142919
BASE COUNT	58266 a	41853 c	40587 g
vector	59992 t		

## ORIGIN

	Query Match	15.7%;	Score 218;	DB 2;	Length 200698;
	Best Local Similarity	57.5%;	Pred. No. 6.7e-51;		
	Matches 392;	Conservative 0;	Mismatches 290;	Indels 0;	Gaps 0;
QY	605	ATGTGAGCCTGATCTCTGCAGAGGTTCCCGTGGGTAGAGTTGTGCTAATGGTAATTTCGCC	664		
Db	119410	AGGTGAGATTTGACAGTGTCTTTCTCTCTCTTAGGGCTTCATCATGACCTTGTCCTCC	119469		
QY	665	TGGTATCTGTCACTATAGGGGCCACCCTTTGCAATATATTTGGCTATCCAGATCAAGTAGC	724		
Db	119470	TGCTGTCTATTGTCTATGGCGCCCTTACGTTGCACAATCTTAGCCCATCAAATCAAGTATG	119529		
QY	725	ATGACTACAGAATTCGCCTTGGGCCACTAGAAGTCTCTGCATCACCATCTGGCGGACAT	784		
Db	119530	ATGAGTATGAGGTCAAAAGTCMAACCCCCTGGCTTATGTCTGTATCTTCTCTCGAAGAGCT	119589		
QY	785	TGAGATACACTTCCCGCCTCTCGATCTTCGTGGTCTCTCTCAGCCCACTTTGAAATTGAAGG	844		
Db	119590	TTGAGATTGCCACCGAGTCATCGTCCGGTCCCTTTTACCTCTGCTCCTGAAGATCTGGG	119649		
QY	845	CTGTGCCCTTCTTAGTGTCTCACTTCCTTGATCATCTCTTTTGAGCCCTTGGATTAAGTTCT	904		
Db	119650	TGGTGGCAGTCATACTCGTCAACTCTTTCAGCTTCTCTTATATCCCTGGATCGTTTTCT	119709		
QY	905	GGAGAAGTGGTGGCCCCAGATGCCCAATAACATTGAGAAAAAATCTGACGCGGTGGCACTC	964		
Db	119710	GGTGCAGTGGCTCCCCATTCCTCGAGAACATAGAGAAGGCCCTAAGTAGGGTGGGTACCA	119769		
QY	965	TGGTGGTCTTGATTTTCAGTCAACATCTCTATGTGGCATCAACTTCTCTTGTGGTCAG	1024		
Db	119770	CCATCGTCTCTGCTTCTCTCACTTACTCTATGCGGGCATCAACATGTTCTGCTGTGUNG	119829		
QY	1025	CTTTGCAGTTTGAGTTTGGCAGACAGAGATCTCGTCGACAAAAGGGCAGAACTGGGGACATA	1084		
Db	119830	CTGTTACAGCTGAAAAATCGACAAATCGGAGCTCATCAGCAAGTCCCAGAACTGGTACCGCC	119889		
QY	1085	TGGCCCTGCACATATAGTGTGAGGTTGGTAGAATATGATCATGGTCTTGGTTTTAAGT	1144		
Db	119890	TGCTCATTTATACATGACGAGATTTCATCGAAGACTCCGTCCTCTGCTCTGTGGTATT	119949		
QY	1145	TCCTTTGAGTGAAGAGTGTACTTGAATTACTCTCATCTCTGATTGCCPTTGCAGGTCATTTA	1204		
Db	119950	TTTTTAAACATGACATATACATGATATGTGTGGCCCTCTGTTGATCTTTCAGCTGCTTTA	120009		
QY	1205	TTGCTTATCTGAATTCATATGGCTTCATGCTCCCTTTTCTCCAGTACTTGCATCCATTGC	1264		
Db	120010	TTGGGTACTGCATGGCATCTCTTCATGCTTGTTGCTTCTATCAGTTTTTCCACCCCTGCA	120069		
QY	1265	GCTCACTCTTCACCCATAATGT	1286		
Db	120070	AAAAGCTCTTCTCCTCCAGTGT	120091		

[illegible]

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Nguyen,C., Maupin,R., Hawkins,M. and Smith,R.  
 The sequence of Homo sapiens BAC clone RP11-506M9  
 Unpublished  
 3 (bases 1 to 166436)  
 Waterston,R.H.  
 Direct Submission  
 Submitted (04-DEC-1999) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 4 (bases 1 to 166436)  
 Waterston,R.H.  
 Direct Submission  
 Submitted (23-APR-2000) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 5 (bases 1 to 166436)  
 Waterston,R.H.  
 Direct Submission  
 Submitted (24-APR-2000) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 6 (bases 1 to 166436)  
 Waterston,R.H.  
 Direct Submission  
 Submitted (23-APR-2000) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 7 (bases 1 to 166436)  
 Waterston,R.  
 Direct Submission  
 Submitted (30-SEP-2000) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Apr 23, 2000 this sequence version replaced gi:5524293.  
 ----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc>  
 Contact: [sapiens@watson.wustl.edu](mailto:sapiens@watson.wustl.edu)  
 ----- Summary Statistics  
 -----  
 Center project name: H\_NH0506M09  
 -----

**NOTICE:** This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

**MAPPING INFORMATION:**

**Materials and Methods:** The position of this clone was established as part of a collaboration between the Human Chromosome Y Mapping Project (Tomoko Kawaguchi, Helen Skaletsky, Laura G. Brown, Steve Rozen, and David C. Page at the Whitehead Institute for Biomedical Research, Cambridge MA) and the Washington University Genome Sequencing Center, St. Louis MO.

**SOURCE INFORMATION:**

The RPE1-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P. Y., Zhao, B., Frenken, E., Iateno, M., Carabese, J. J. and de Jong, P. J. (1998). An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>) (VECTOR: pBACe3.6)

## NEIGHBORING SEQUENCE INFORMATION:

This clone sequenced to the left is RP11-221K4. Actual start of this clone is at base position 1 of RP11-506M9; actual end is at base position 166436 of RP11-506M9.

## FEATURES

## Location/Qualifiers

```

source
1..166436
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="Y"
/map="Y"
/clone="RP11-506M9"
/clone_lib="RPCI-11"
195..923
/rpt_family="L1"
924..1077
/rpt_family="L1"
1080..1830
/rpt_family="L1"
1879..1918
/rpt_family="L1"
1905..3257
/rpt_family="L1"
3280..3393
/rpt_family="Alu"
3618..3694
/rpt_family="MIR"
4250..4511
/rpt_family="Alu"
4512..5804
/rpt_family="L1"
5804..7682
/rpt_family="L1"
8188..8294
/rpt_family="L2"
9429..9499
/rpt_family="L2"
9826..9872
/rpt_family="(A)n"
9962..10035
/rpt_family="L2"
10596..10684
/rpt_family="GA-rich"
12836..13209
/rpt_family="MaLR"
14667..14689
/rpt_family="AT-rich"
16692..16712
/rpt_family="AT-rich"
17516..17584
/rpt_family="MERS3"
17590..17667
/rpt_family="MERS3"
17814..18112
/rpt_family="Alu"
19421..19650
/rpt_family="L1"
19689..19813
/rpt_family="Retroviral"
19815..20107
/rpt_family="Retroviral"
20410..20861
/rpt_family="Retroviral"
21083..25178
/rpt_family="Retroviral"
25196..27960
/rpt_family="Retroviral"
28006..28520
/rpt_family="Retroviral"
28828..28905
/rpt_family="L2"
29157..29395
/rpt_family="Alu"
29396..29438

```

```

repeat_region
30965..30996
/rpt_family="TAA)n"
repeat_region
31776..32149
/rpt_family="AT-rich"
repeat_region
33205..33489
/rpt_family="MaLR"
repeat_region
35802..35832
/rpt_family="Alu"
repeat_region
36311..36389
/rpt_family="AT-rich"
repeat_region
36757..36827
/rpt_family="(TTAAA)n"
repeat_region
38147..38551
/rpt_family="MERS3"
repeat_region
38552..38576
/rpt_family="Retroviral"
repeat_region
38577..38633
/rpt_family="(TG)n"
repeat_region
38634..39399
/rpt_family="Retroviral"
repeat_region
39400..39459
/rpt_family="Retroviral"
repeat_region
39460..40129
/rpt_family="MERI_type"
repeat_region
40121..41902
/rpt_family="Retroviral"
repeat_region
41931..42280
/rpt_family="Retroviral"
repeat_region
42281..42759
/rpt_family="Retroviral"
repeat_region
42961..43305
/rpt_family="Retroviral"
repeat_region
43410..43703
/rpt_family="L1"
repeat_region
44044..44848
/rpt_family="Alu"
repeat_region
44853..45060
/rpt_family="L1"

```

Query Match 15.6%; Score 217.2; DB 9; Length 166436;

Best Local Similarity 62.3%; Pred. No. 1.1e-50;

Matches 391; Conservative 0; Mismatches 233; Indels 4; Gaps 3;

```

Qy 642 AGTTGTCTAATGGTATTTCCCTGGTATCTGTACCTATGGGCCACCCCTTTGCAATAT 701
Db 114404 AGCAATGCTGATGACATGTTCCCTCTTACCACCTGTGTATGGGCCATTCACCTGCAATAC 114345
Qy 702 GTTGGCTATCCAGATCAAGTACGATGACTACAAGATTCGCTTGGGCCACTAGAAAGTCT 761
Db 114344 ACTGGCCATCCATACAGCAGCAGCTTCTCGACATTAAGCTACAGCGATAAAATTCAT 114285
Qy 762 CTGCATCACCATCTGGCGGACATTTGGAGATCATTTCCGCCCTCCTGTGATTTGGTGCTCTT 821
Db 114284 CTG--TCATGATACGGTGTCTCTTGGTGATTTATCTCAGCTAGTAGCTCTGGAATTTT 114227
Qy 822 CTGAGCCACTTTGAAATGAAGCTGTGCCCTTCTAGTGTCTCAACTTCCTGATCATCT 881
Db 114226 TCCTGCATCTCTGAACAGAGAGCCTACCCCTTACTATTATCATATATTTTGTATTAT 114167
Qy 882 CTTTGAGCCCTGGATTAAAGTTCTGGAGAAGTGGTGGCCAGATGCCCAATAACATTGAGAA 941
Db 114166 GTTGGCACCATGGCTGGAGTTTGGAAAAGTGGAGCTCATCTTATATAACAACAGAAAA 114107
Qy 942 AAACCTTCAGCCGGTGGCGCACTCTGTGGTCTCTGATTTCACTACCATTCCTCTATGCTGG 1001
Db 114106 TAATTCAGCATGGTGGGTAC-ATACGATGCTTACCTTAATCATGTCTACTATATGCTGC 114048
Qy 1002 CATCAACTTCTCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1061
Db 114047 TATCAACTTCTCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 113988
Qy 1062 CAAAGGCGAGAACTGGGGACATATGGGCCCTGCCTAGTGTAGGTTGGTGGTGGTGGTGGTGG 1121

```

```

from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
The position of this clone was established as part of a
collaboration between the Human Chromosome Y Mapping Project
(Tomoko Kawaguchi, Helen Skaletsky, Laura G. Brown, Steve Rozen,
and David C. Page at the Whitehead Institute for Biomedical
Research, Cambridge MA) and the Washington University Genome
Sequencing Center, St. Louis MO.

SOURCE INFORMATION:
The RPL11 human BAC library was made from the blood of one male
donor, as described by Oseogawa, K., Woon, P.Y., Zhao, B., Frengen, E.,
Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at the Roswell Park Cancer Institute
(http://bacpac.med.buffalo.edu)

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-79J10; the clone sequenced
to the right is RP11-45K3, 200 bp overlap. Actual start of this
clone is at base position 1 of RP11-100J21; actual end is at base
position 66964 of RP11-245K4.

Location/Qualifiers
1. .161879
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="Y"
/map="Y"
/clone="RP11-100J21"
/clone.lib="RPCI-11"
1. .1185
/rpt_family="L1"
1183. .1513
/rpt_family="L1"
1679. .1842
/rpt_family="MER1_type"
2125. .3135
/rpt_family="ERV1"
3116. .3412
/rpt_family="ERV1"
3450. .3866
/rpt_family="ERV1"
3867. .6977
/rpt_family="ERV1"
6986. .7295
/rpt_family="ERV1"
7297. .7625
/rpt_family="ERV1"
9507. .9570
/rpt_family="MER1_type"
11143. .11817
/rpt_family="L1"
12324. .12843
/rpt_family="L1"
12874. .12922
/rpt_family="ERV1"
12885. .12929
/rpt_family="ERV1"
13051. .13264
/rpt_family="L1"
13381. .13619
/rpt_family="L1"
13628. .14306
/rpt_family="L1"
15532. .15662
/rpt_family="ERV1"
15740. .16044
/rpt_family="ERV1"
16045. .16136
region

```

repeat\_region /rpt\_family="ERV1"  
16178. .16239 /rpt\_family="ERV1"  
16728. .17201 /rpt\_family="ERV1"  
17188. .17253 /rpt\_family="ERV1"  
17243. .17602 /rpt\_family="ERV1"  
17746. .18064 /rpt\_family="ERV1"  
18085. .18235 /rpt\_family="L1"  
18303. .18876 /rpt\_family="ERVK"  
18927. .19413 /rpt\_family="ERVK"  
19419. .19628 /rpt\_family="ERVK"  
21508. .21759 /rpt\_family="L1"  
22725. .22868 /rpt\_family="Alu"  
22868. .23335 /rpt\_family="ERV1"  
23350. .24076 /rpt\_family="ERV1"  
24273. .24690 /rpt\_family="ERV1"  
24702. .24823 /rpt\_family="ERV1"  
24886. .25175 /rpt\_family="Alu"  
25186. .25543 /rpt\_family="ERV1"  
27654. .27697 /rpt\_family="ERV1"  
28186. .28421 /rpt\_family="ERV1"  
28440. .28804 /rpt\_family="ERV1"  
28773. .29267 /rpt\_family="ERV1"  
29270. .31939 /rpt\_family="ERV1"  
31933. .32211 /rpt\_family="ERV1"  
32225. .32546 /rpt\_family="Alu"  
32582. .33624 /rpt\_family="ERV1"  
33740. .34049 /rpt\_family="ERV1"  
34050. .34149 /rpt\_family="ERV1"  
34155. .34617 /rpt\_family="ERV1"  
34624. .37556 /rpt\_family="ERV1"  
37676. .37979 /rpt\_family="ERV1"  
38072. .38379 /rpt\_family="ERV1"  
38380. .39312 /rpt\_family="ERV1"  
39318. .40245 /rpt\_family="ERV1"  
40372. .40443 /rpt\_family="ERV1"  
40486. .41009 /rpt\_family="ERV1"

Query Match

15.5%; Score 215.6; DB 9; Length 161879;

Best Local Similarity 62.1%; Pred. No. 3.2e-50;  
Matches 390; Conservative 0; Mismatches 234; Indels 4; Gaps 3;  
Qy 642 AGTTGTGCTAATGGTATTTTCCTCGGTATCTGTCACTATATGGGGCCACCCCTTTGCAATAT 701  
Db 160312 AGCATTTGCTGATGACATGTTCTTCTTACCACTTGTGTATGGGCCATTTCACTGCAATAC 160371  
Qy 702 GTTGGCTATCCAGATCAAGTACGATGACTACAAGATTCCCTTGGGCCACTAGAATCCT 761  
Db 160372 ACTGGCCATCCATACCAGCAAGCACTTTCCTGACATTAAGCTACAGGCAATAAATTCAT 160431  
Qy 762 CTGCATCAACCATCTGGCGGACATTTGGAGATCACTTCCCGCCTCCTGATTTCTGGTCTCTT 821  
Db 160432 CTG--TCATGATACGGTGTTCCTTGGTGATTAATCTCACGTGTAGTACTCTGGAATTTT 160489  
Qy 822 CTCAGCCACTTTGAAATTTGAAGGCTGTGCCCTTCTCTAGTGTCAACTTCTCTGATCATCCT 881  
Db 160490 TCCTGCATCTCTGAAACACAGAGAGCCTTACCACTTACTATTAATCATATATTTTGTATT 160549  
Qy 882 CTTTGAGCCCTGGATTAAAGTTTCTGGAGAAAGTGTCGCCAGATGCCCAATAACATTGAGAA 941  
Db 160550 GTTGCACCATGGCTGGAGTTTGGAAAGTGGAGCTCATCTTCATAACACACAGAAAA 160609  
Qy 942 AAACCTTCAGCCGGTGGCAGCTCTGGTGGTCTCTGATTTTCAGTCAACCATCTCTATCTCTGG 1001  
Db 160610 TAAATTCAGCATGGTGGGTAC-ATACGGATGCTTACCTTAATCATGCTACTATATCTGTC 160668  
Qy 1002 CATCAACTTCTCTGCTGGTGCAGCTTTTCGAGTTGAGTTGGCAGACAGAGATCTCGTCGA 1061  
Db 160669 TATCAACTTCTCTGGCTGGTCAGCACTGAAACTGCAGCTCTCAAAATGAGGAAATTAATGA 160728  
Qy 1062 CAAGGGCAGAACTGGGGCAGATATGGGCTGCACATATAGTGTGAGGTTGGTAGAGAAATGT 1121  
Db 160729 CAAGAGACTGAGGTGGGC-CATAGAATCCTACACTACAGCTTCTAGTTTTTTAGAAAATGT 160787  
Qy 1122 GATCATGCTCTTGGTTTTTAAAGTTCTTTGGAGTGAAGTGTGTACTGAATTAATCTGTCAATTC 1181  
Db 160788 GATAATAATATTGATATTATTTATGTTTCGTTGGAGGAAATTTTCACCTGAAGTCTTTGTGACTC 160847  
Qy 1182 CTGATTTGCCCTTGCAAGCTCATTTATGCTTTATCTGATTTTCCATTTGCGTTCATGCTCCTTTT 1241  
Db 160848 ATTAATTGCCATGTAGTTTCATCAAACTAACTAATTAGCCATTGGCTTTATGCTCCTCTTT 160907  
Qy 1242 CTTCCAGTACTTGCATCCATTGGGCTCA 1269  
Db 160908 CTGTCAGTATTTCACCAAGGTGGTCA 160935

Search completed: March 30, 2003, 05:28:41  
Job time : 3999.31 secs